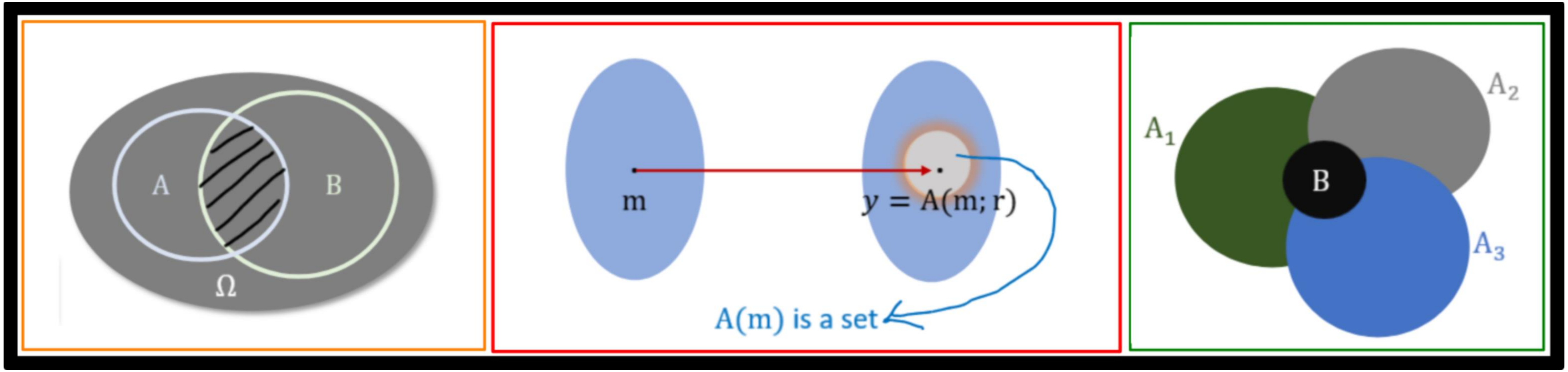


CS409m: Introduction to Cryptography

Lecture 02 (01/Aug/25)

Instructor: Chethan Kamath

(Slides: Sruthi Sekar)

Announcements

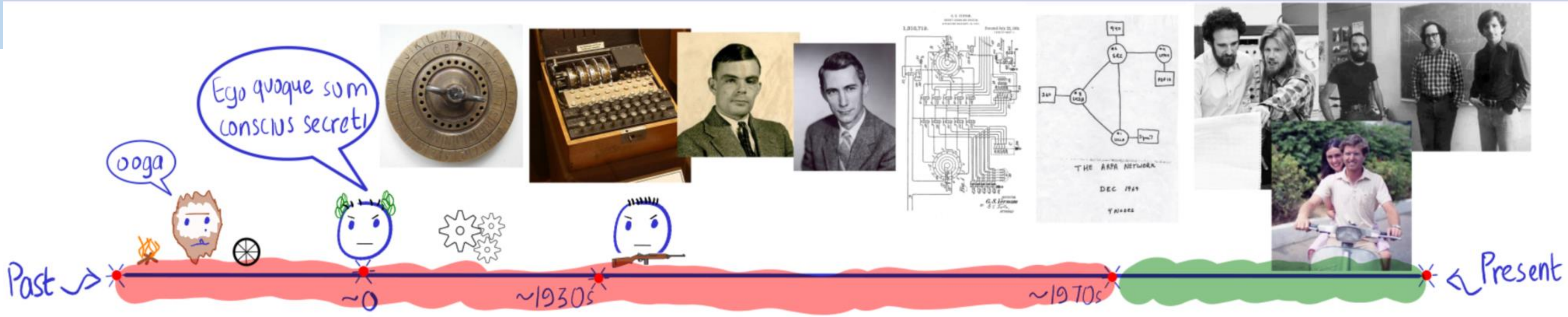
Administrivia...

- Venue change: CC103 to CC101
- Join WhatsApp group (link shared on Moodle)
- TA session: Fridays, 19:30-21:00, CC101

Coursework

- Hands-on Exercise 0 uploaded on Moodle
- Assignment 1 will be uploaded today on webpage/Moodle
  - Ungraded, but you will get some questions in the quiz!

Recall from Last Lecture



- Classical vs modern cryptography
- Guiding principles for modern cryptography:
  - Formally define threat model M: identify security goal and adversary's capabilities
  - Construct scheme
  - Provide rigorous security proof
    - Rely on precise, well-studied assumption
- Classical ciphers: shift cipher, substitution cipher, polyalphabetic shift cipher
- Saw why these are considered insecure by modern standards
  - The ciphertext leaks some information about the message

Plans for Today's Lecture

1. Randomized Algorithms

2. Basic Theory
   - Definition and Background
   - Law of Total Probability, Conditional Probability, Expectation and Variance

3. Concentration Inequalities
   - Union Bound, Markov’s inequality, Chebyshev’s inequality, Chernoff bounds

4. Misc

Motivation: Randomized Algorithms

Deterministic Algorithm

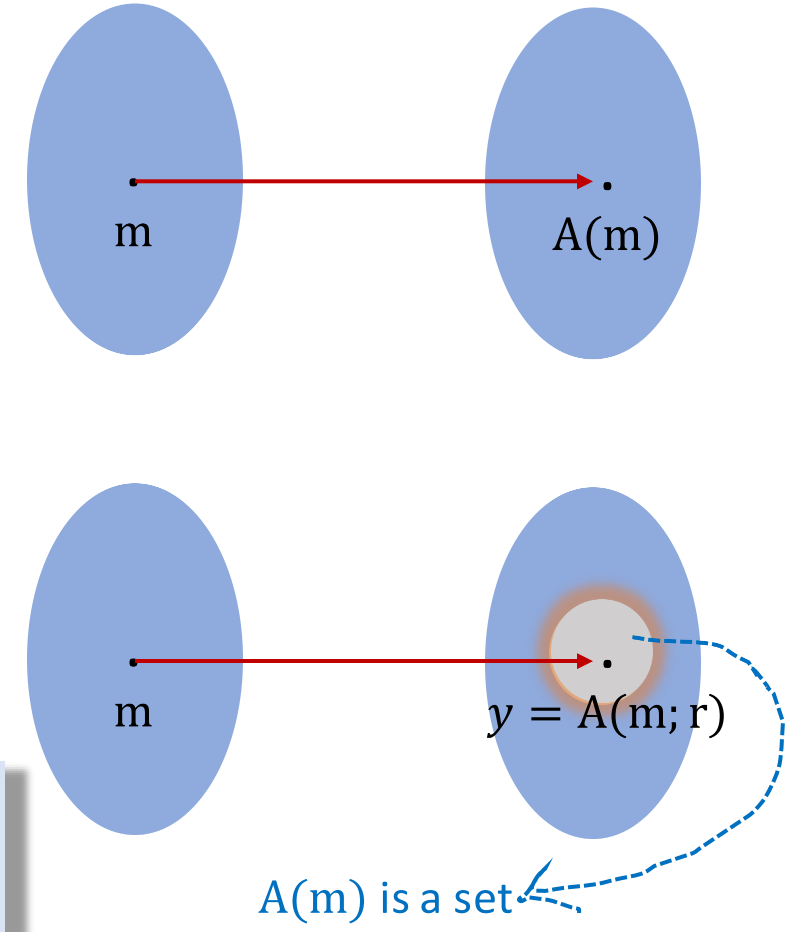
For each input m, A outputs a fixed value y from the codomain
$\mathrm{y} \coloneqq \mathrm{A}(\mathrm{m})$

Randomized Algorithm

For each input m, A additionally uses a randomly picked
$\mathrm{r} \in_R \{0,1\}^{\mathrm{n}}$ and outputs $\mathrm{y} \coloneqq \mathrm{A}(\mathrm{m};\mathrm{r})$.
($\mathrm{A}(\mathrm{m})$ is now a set instead of a single value)

Why?

- Randomized algorithms can be faster than deterministic ones
- In cryptography, randomization will be necessary for security
    - Key generation and encryption will be randomized



BASIC THEORY

Definition and Background

Ω: finite set of possible outcomes, called sample space of the experiment

Example: Coin flip
$\Omega = \{H, T\}$

Definition 1

A probability distribution over Ω is a function $P\colon \Omega \to \mathbb{R}_{\geq 0}$ such that $\Sigma_{x \in \Omega} P(x) = 1$.
(sum of all p_i's over all possible outcomes in Ω is 1)

Example: Coin flip
$P(H) = P(T) = 1/2$

Definition 2

An event is any subset $A \subseteq \Omega$. The probability of an event A is $\Pr_P[A] = \Sigma_{x \in A} P(x)$.

Definition 3

Two events A and B are said to be independent if $\Pr[A \cap B] = \Pr[A] . \Pr[B]$.

Examples

Example 1: Tossing a coin n times

$$\Omega = \{(\omega_1, \omega_2, \ldots \omega_n): \omega_i \in \{H, T\}\}$$

The event that the first flip is heads is represented as:

$$A_{1,H} = \{(H, \omega_2, \ldots \omega_n): \omega_i \in \{H, T\}$$

The event that the first flip is tails is represented as:

$$A_{1,T} = \{(T, \omega_2, \ldots \omega_n): \omega_i \in \{H, T\}$$

Let $A_{i,H}$ be the event that the i-th flip is H, and similarly define $A_{i,T}$. Suppose coin is fair ($P[H] = P[T]$) and each toss is independent, then we have: for a fixed $\omega_1, \omega_2, \ldots \omega_n$

$$\begin{aligned} P((\omega_1, \omega_2, \ldots \omega_n)) &= \Pr[A_{1,\omega_1} \cap \cdots \cap A_{n,\omega_n}] \\ &= \Pr[A_{1,\omega_1}] \ldots \Pr[A_{n,\omega_n}] && \text{(by independence)} \\ &= \frac{1}{2} \cdots \frac{1}{2} = \frac{1}{2^n} && \text{(by fairness)} \end{aligned}$$

Definition and Background

Definition 4

A (real-valued) random variable is a function $\mathrm{X}\colon\ \Omega\ \rightarrow \mathbb{R}$. In Example 1, the number of heads is a random variable represented by the function $\mathrm{X}\big((\omega_1, \omega_2, \dots \omega_n)\big) = \Sigma_{i=1}^{n} \omega_i$.

Definition 5

Two discrete real-valued random variables, X and Y, are said to be independent if
$$\Pr[X = x, Y = y] = \Pr[X = x].\Pr[Y = y], \quad \forall\, x, y \in \mathbb{R}.$$

The random variables $X_1, \dots X_n$ are said to be jointly independent if
$$\Pr[X_1 = x_1, \dots, X_n = x_n] = \Pi_{i=1}^{n} \Pr[X_i = x_i]\,, \forall\, x_1, \dots, x_n \in \mathbb{R}$$

Examples

Example 2: Jointly independent random variables

$$X_i = \begin{cases} 1, \text{if } i^{th} \text{coin lands H} \\ 0, \text{if } i^{th} \text{coin lands T} \end{cases}$$

The random variables $X_1, \ldots X_n$ are jointly independent.

Example 3: Pairwise Independent but not jointly independent random variables

$\Omega = \{(\omega_1, \omega_2): \omega_i \in \{0,1\}\}$. Consider uniform distribution P over Ω.

$X(\omega_1, \omega_2) = \omega_1$ (outcome of first flip)

$Y(\omega_1, \omega_2) = \omega_2$ (outcome of second flip)

$Z(\omega_1, \omega_2) = \omega_1 \oplus \omega_2$ (XOR of both flips)

$\forall\, x, y, z \in \{0,1\}, \Pr[X = x] = \Pr[Y = y] = \Pr[Z = z] = 1/2$, and

$\Pr[X = x, Y = y] = \Pr[X = x, Z = z] = \Pr[Y = y, Z = z] = 1/4 \Longrightarrow$ pairwise independence

But, $\Pr[X = 0, Y = 0, Z = 1] = 0 \neq 1/8 = \Pr[X = 0] . \Pr[Y = 0] . \Pr[Z = 1] \Longrightarrow$ not jointly independent

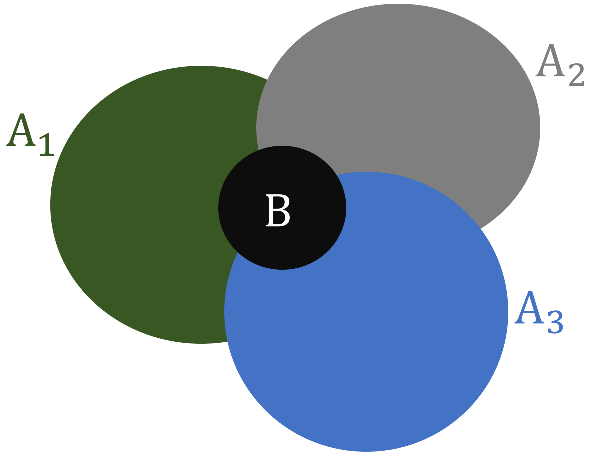
Law of Total Probability

If we have events $A_1, A_2, \ldots A_n$ that partition the sample space Ω (i.e., Ω is a disjoint union of these sets), and let B be any event, then

$$\Pr[B] = \Sigma_{i=1}^{n} \Pr[B \cap A_i]$$

EXAMPLE:

$$\Omega = A_1 \sqcup A_2 \sqcup A_3$$



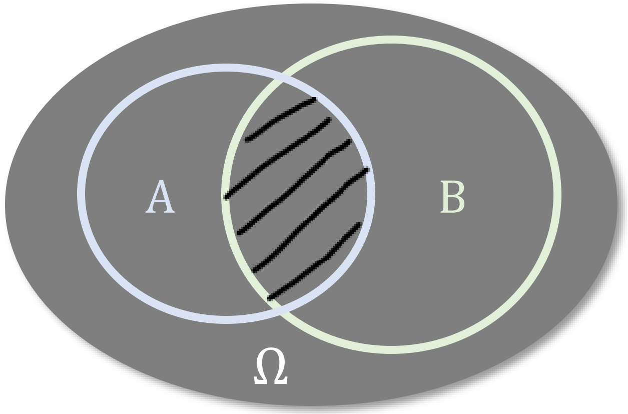
Conditional Probability

Definition 6

Probability of an event A conditioned on event B (with $\Pr[B] \neq 0$)) is defined as

$$\Pr[A|B] = \frac{\Pr[A \cap B]}{\Pr[B]}$$

Conditioning on B means you are now considering B as the sample space instead of Ω.



Bayes' Rule

$$\Pr[A|B] = \frac{\Pr[B|A]\Pr[A]}{\Pr[B]}$$

Chain Rule

$$\Pr[A_1 \cap A_2 \cap \cdots \cap A_n] = \Pr[A_1]\Pr[A_2|A_1] \ldots \Pr[A_n|A_1 \cap A_2 \cap \cdots \cap A_{n-1}]$$

Expectation

Definition 7

For a discrete real-valued random variable X taking possible values $x_1, x_2, \ldots, x_n \in \mathbb{R}$, the expectation is defined as

$$\mathbb{E}[X] = \sum_{i=1}^{n} \Pr[X = x_i] \cdot x_i$$

"average" value that X will take

Example 4: Expectation of a coin toss

$$X = \begin{cases} 1, \text{if a fair coin lands H} \\ 0, \text{otherwise} \end{cases}$$

$$\mathbb{E}[X] = \frac{1}{2} \cdot 1 + \frac{1}{2} \cdot 0 = \frac{1}{2}$$

Properties of Expectation

Linearity of Expectation

Given random variables $X_1, \dots, X_n$ and $X = \Sigma_{i=1}^{n} X_i$, we have

$$\mathbb{E}[X] = \mathbb{E}[\Sigma_{i=1}^{n} X_i] = \sum_{i=1}^{n} \mathbb{E}[X_i]$$

**Holds even if the X_i's are not independent!

Example 4: Coin tosses

$X_i = \begin{cases} 1, \text{if the } i^{th} \text{ fair coin lands H} \\ 0, \text{otherwise.} \end{cases}$ and let $X = \Sigma_{i=1}^{n} X_i$

$$\mathbb{E}[X] = \mathbb{E}[\Sigma_{i=1}^{n} X_i] = \sum_{i=1}^{n} \mathbb{E}[X_i] = \frac{n}{2}$$

Multiplicativity of Expectation under independence

For independent random variables X and Y

$$\mathbb{E}[XY] = \mathbb{E}[X].\,\mathbb{E}[Y]$$

HW: Prove this by expanding the right side of the equation and using the independence.

Variance

Definition 8

For a discrete real-valued random variable X the variance is defined as

$$\mathbf{Var}[X] = \mathbb{E}[(X - \mathbb{E}[X])^2] = \mathbb{E}[X^2] - \mathbb{E}[X]^2$$

Variance measures how far the random variable deviates from its expectation

Linearity of variance under pairwise independence

Given random variables $X_1, \ldots, X_n$ that are pairwise independent

$$\mathbf{Var}[\Sigma_{i=1}^{n} X_i] = \sum_{i=1}^{n} \mathbf{Var}[X_i]$$

HW: Prove this using properties of expectation!

CONCENTRATION INEQUALITIES

Union Bound

For events $A_1, A_2, \dots A_n \subseteq \Omega$,

$$\Pr\left[\bigcup_{i=1}^{n} A_i\right] \leq \sum_{i=1}^{n} \Pr[A_i]$$

Proof

For $A_1, A_2 \subseteq \Omega$

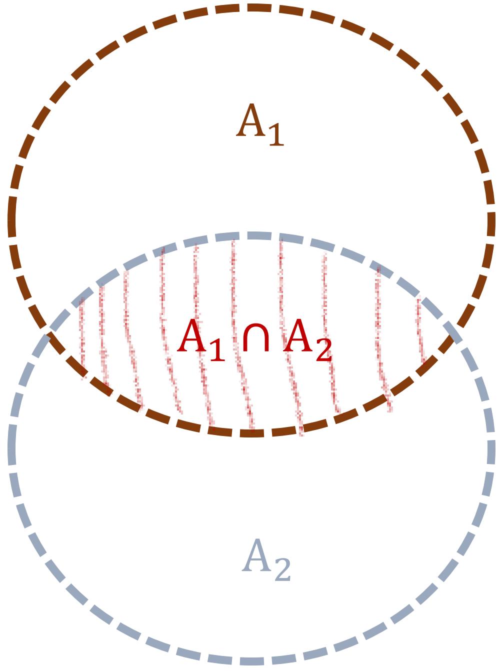
$$\Pr[A_1 \cup A_2] = \Pr[A_1] + \Pr[A_2] - \Pr[A_1 \cap A_2]$$

In particular

$$\Rightarrow \Pr[A_1 \cup A_2] \leq \Pr[A_1] + \Pr[A_2]$$

$\therefore$ By induction, for $A_1, \dots, A_n \subseteq \Omega$

$$\Pr\left[\bigcup_{i=1}^{n} A_i\right] \leq \sum_{i=1}^{n} \Pr[A_i] \quad \blacksquare$$



Inclusion-exclusion Principle

Markov's Inequality

For a discrete random variable X taking non-negative values in the set S, and for any $\alpha > 0$,

$$\Pr[X \geq \alpha] \leq \frac{\mathbb{E}[X]}{\alpha}$$

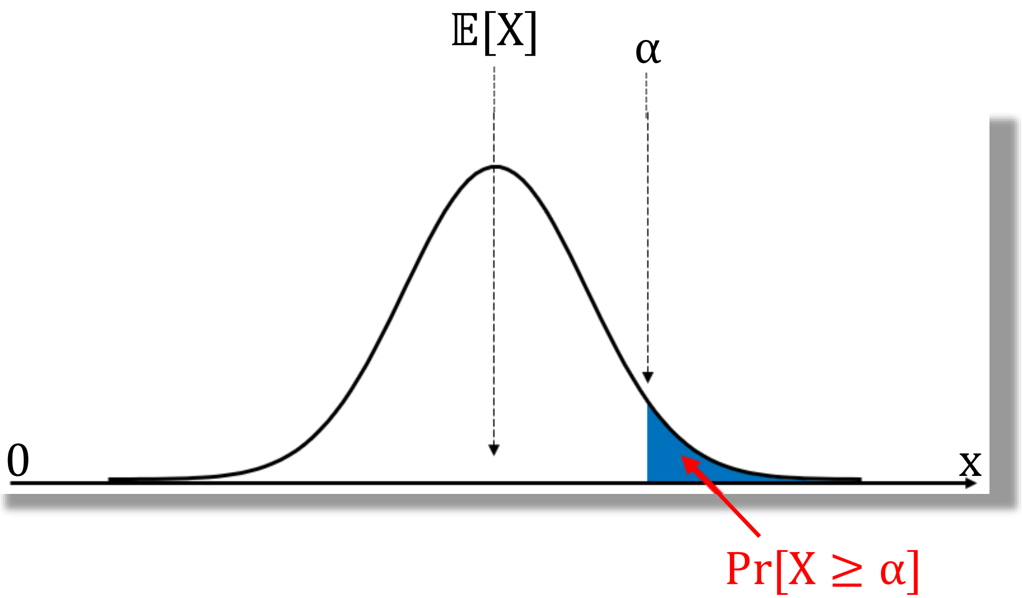
Proof:

$$\mathbb{E}[X] = \sum_{x \in S} x \cdot \Pr[X = x]$$

$$= \sum_{\substack{x \in S \\ x < \alpha}} x \cdot \Pr[X = x] + \sum_{\substack{x \in S \\ x \geq \alpha}} x \cdot \Pr[X = x]$$

$$\geq \sum_{\substack{x \in S \\ x \geq \alpha}} x \cdot \Pr[X = x]$$

$$\geq \sum_{\substack{x \in S \\ x \geq \alpha}} \alpha \cdot \Pr[X = x] = \alpha \cdot \Pr[X \geq \alpha]$$



$\blacksquare$

Chebyshev's Inequality

For a discrete random variable X with variance $\sigma^2 > 0$, and for all real $\alpha > 0$

$$\Pr[|X - \mathbb{E}[X]| \geq \alpha] \leq \frac{\sigma^2}{\alpha^2}$$

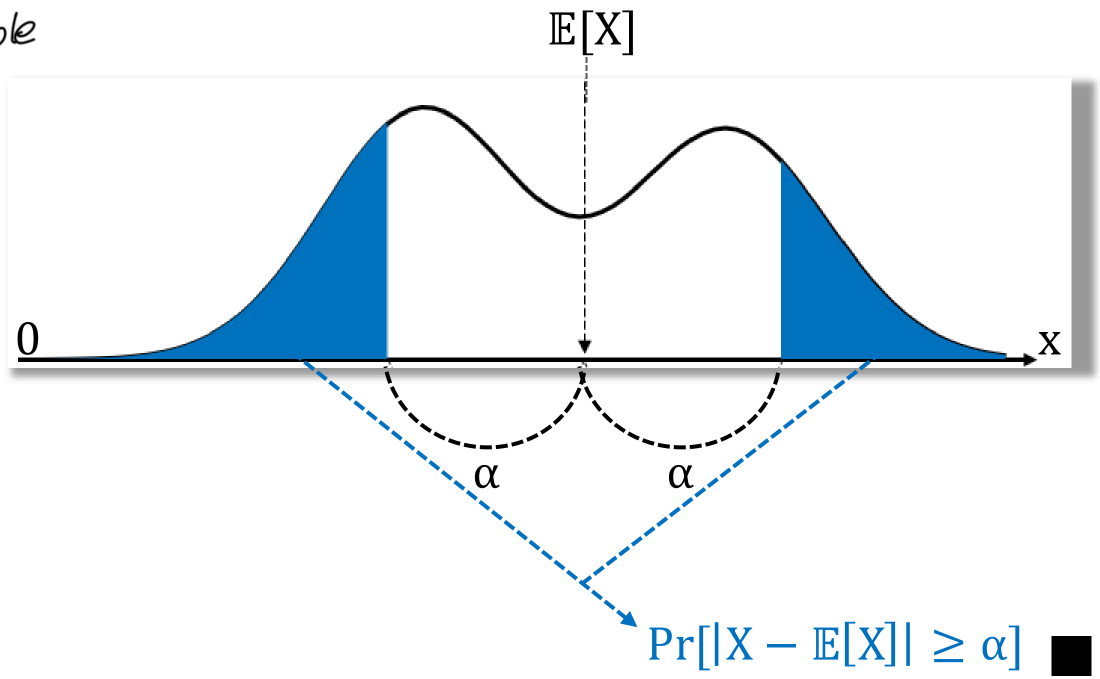
Proof:

$\because$ $(X-\mathbb{E}[X])^2$ is non-negative random variable

by Markov's inequality

$$\Pr[(X-\mathbb{E}[X])^2 \geq \alpha^2] \leq \frac{\mathbb{E}[(X-\mathbb{E}[X])^2]}{\alpha^2}$$

$$\Rightarrow \Pr[|X-\mathbb{E}[X]| \geq \alpha] \leq \frac{\sigma^2}{\alpha^2}$$



■

Chernoff Bounds

Suppose $X_1, \ldots, X_n$ are independent random variables taking values in $\{0,1\}$. For any $\alpha > 0$

- $\Pr[\Sigma_{i=1}^{n} X_i > (1+\alpha)\mathbb{E}[X]] < e^{-\alpha^2 \mathbb{E}[X]/3}$, for $0 < \alpha \leq 1$
- $\Pr[\Sigma_{i=1}^{n} X_i > (1+\alpha)\mathbb{E}[X]] < e^{-\alpha \mathbb{E}[X]/3}$, for $\alpha > 1$
- $\Pr[\Sigma_{i=1}^{n} X_i < (1-\alpha)\mathbb{E}[X]] < e^{-\alpha^2 \mathbb{E}[X]/2}$, for $0 < \alpha < 1$

Tighter bounds making use of joint independence!

MISC

(We'll need these later.)

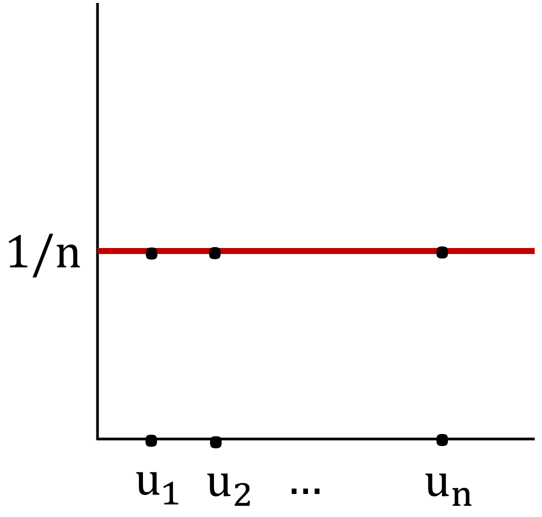
Uniform Random Variable

Definition 9

X is a uniform random variable over a set S, if,

$$\forall\, u \in S, \Pr[X = u] = 1/|S|$$

(denote the distribution by U_S, and picking an element under uniform distribution by $u \in_R S$)



$S = \{u_1, u_2, \dots, u_n\}$

XOR and its Properties

XOR of two strings in $\{0,1\}^n$ is their bit-wise addition modulo 2.

X	Y	$X \oplus Y$
0	0	0
0	1	1
1	0	1
1	1	0

XOR of bits

$$\begin{array}{r} 0\,1\,1\,0\,1\,1\,1 \\ \oplus\; 1\,0\,1\,1\,0\,1\,0 \\ \hline 1\,1\,0\,1\,1\,0\,1 \end{array}$$

Example

Theorem 1

Let Y be a random variable over $\{0,1\}^n$ and X be an independent uniform random variable over $\{0,1\}^n$.
Then $Z = X \oplus Y$ is a uniform random variable on $\{0,1\}^n$.

HW: Prove this
(Hint: use induction)

Birthday Paradox

Let q birthdays be $y_1, y_2, \ldots, y_q$ chosen uniformly at random from $\{1,2,\ldots,365\}$

*assuming non-leap year and that birthdays are uniform and independent

<u>Problem</u>: Find minimal q such that $\mathrm{coll}(q,365) = \Pr[\exists\, i \neq j \text{ s.t. } y_i = y_j] \geq 1/2$

HW: Solve for q by using the same argument with unknown q

INTUITION:

23

comparisons with 23 people in the room

$= 22 + 21 + 20 + \cdots + 1$

$= 253$

Pr[None of these 253 comparisons match]

$= \left(\frac{364}{365}\right)^{253} = 49.95\%$

↳ probability one single comparison not matching.

$\Rightarrow$ Pr[At least 1 pair matches] = 50.05%

Birthday Paradox, Generalized

Theorem 2

Let q elements be $y_1, y_2, \ldots, y_q$ be chosen uniformly and independently at random from a set of size N, then

$$\mathrm{coll}(q, N) = \Pr[\exists\, i \neq j \text{ s.t. } y_i = y_j] \leq \frac{q^2}{2N}$$

Assignment Problems

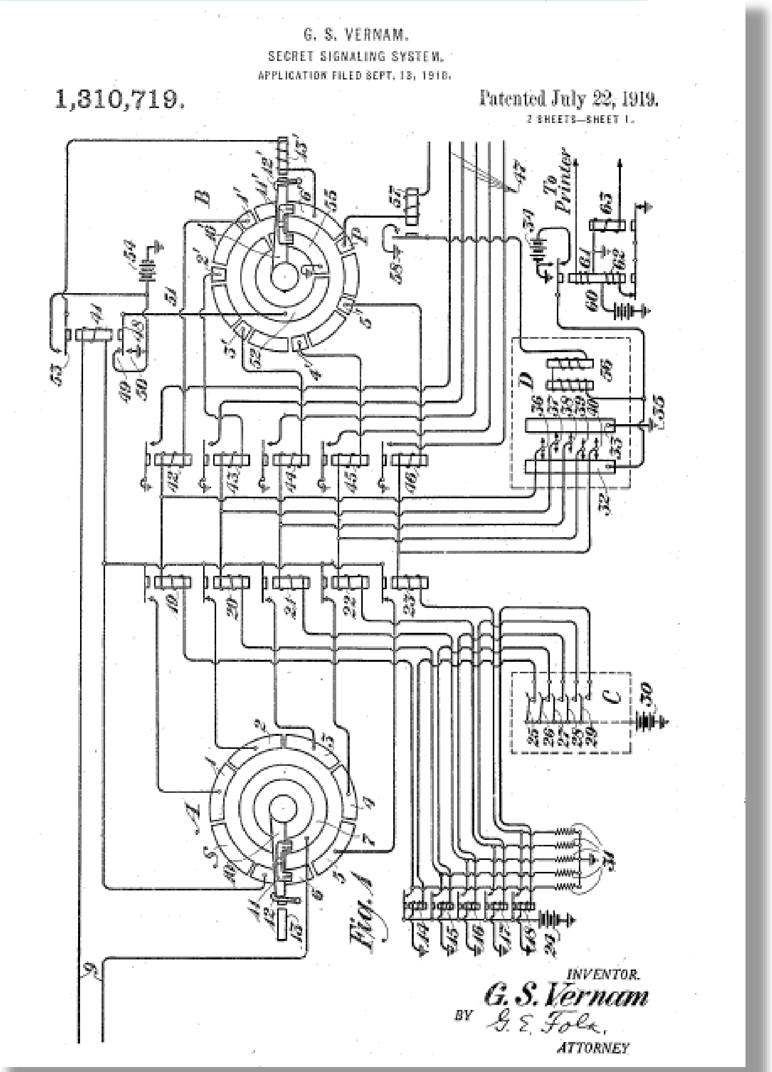
Theorem 3

Let $q \leq \sqrt{2N}$ elements be $y_1, y_2, \ldots, y_q$ be chosen uniformly and independently at random from a set of size N, then

$$\mathrm{coll}(q, N) = \Pr[\exists\, i \neq j \text{ s.t. } y_i = y_j] \geq \frac{q(q-1)}{4N}$$

Next Lecture

- Perfect security for shared-key/symmetric encryption
- Example: one-time pad or Vernam cipher
- Limitations of perfect security and some attacks



Thank you!